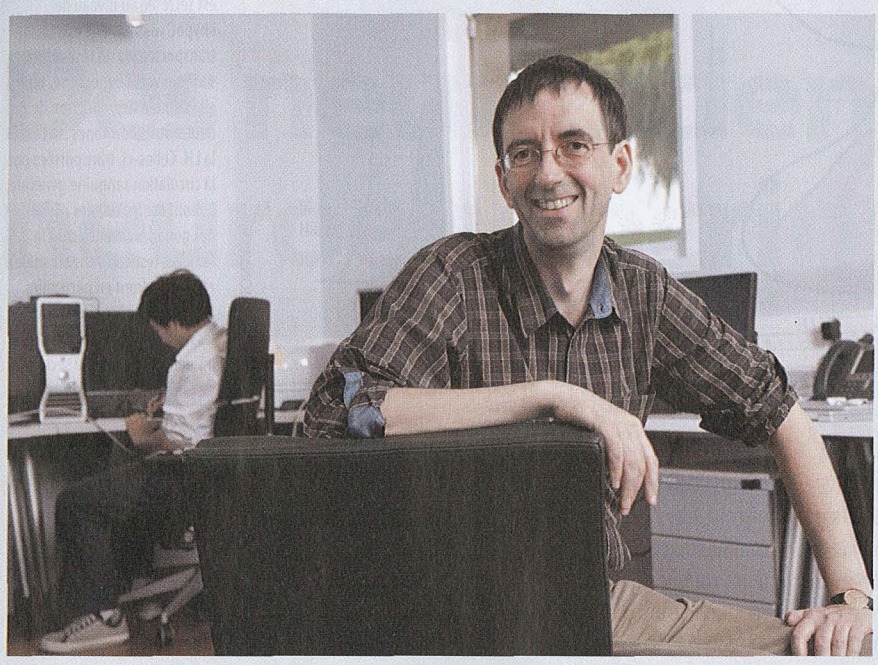
Un laboratoire de haute sécurité en informatique

ENTRETIEN AVEC JEAN-YVES MARION



© INRIA-KAKSONEN

Jean-Yves Marion, professeur à Nancy Université (Institut national polytechnique de Lorraine), est directeur du laboratoire de haute sécurité (LHS). Il est également responsable de l'équipe-projet CARTE (Inria Nancy-Grand Est), spécialisée en virologie informatique.

Dans la seule année 2009, on a dénombré plus de virus que dans toute l'histoire de l'informatique. C'est pour trouver des parades à ces attaques et plus généralement aux divers problèmes de sécurité informatique qu'a été créé le laboratoire nancéen de haute sécurité (LHS).

Les questions de sécurité informatique ne datent pas d'aujourd'hui, loin de là. Pourquoi avoir tant attendu pour créer un laboratoire de haute sécurité (LHS) ?

Jean-Yves Marion : J'avoue que je me suis moi-même posé la question… je n'ai pas la réponse. Toujours est-il que c'est le premier du genre en France (si l'on excepte des laboratoires militaires), de même qu'aucun laboratoire académique de ce type n'existe en Europe. On trouve en revanche un équivalent du notre à l'École polytechnique de Montréal et des structures analogues aux États-Unis. Les recherches et expériences menées au LHS couvrent la virologie informatique, les outils de surveillance des réseaux et la recherche de vulnérabilité, en particulier des équipements*.

Les enjeux économiques de la sécurité informatique sont pourtant énormes…

J.-Y. M. : Bien entendu, et ce à divers niveaux. Il y a d'abord ceux liés aux attaques des PC d'utilisateurs individuels. Il y a ensuite ceux liés au vol et à la destruction d'informations, notamment dans les entreprises. Enfin, le troisième niveau est celui de la cybercriminalité, du cyberterrorisme… Les risques sont en outre d'autant plus grands que les dégâts causés par un code malveillant peuvent désormais sortir du monde virtuel des ordinateurs. La démonstration éclatante en a été faite par le fameux ver informatique *stuxnet*, découvert en juin 2010. Celui-ci était capable de reprogrammer un système industriel comme une centrale nucléaire *via* le réseau informatique. Un rapport récent de l'Organisation européenne pour la sécurité (EOS) évalue à 350 milliards d'euros l'impact global des menaces informatiques en Europe. Quant au budget de l'économie souterraine liée à la sécurité informatique, il atteindrait quelque 100 milliards d'euros.

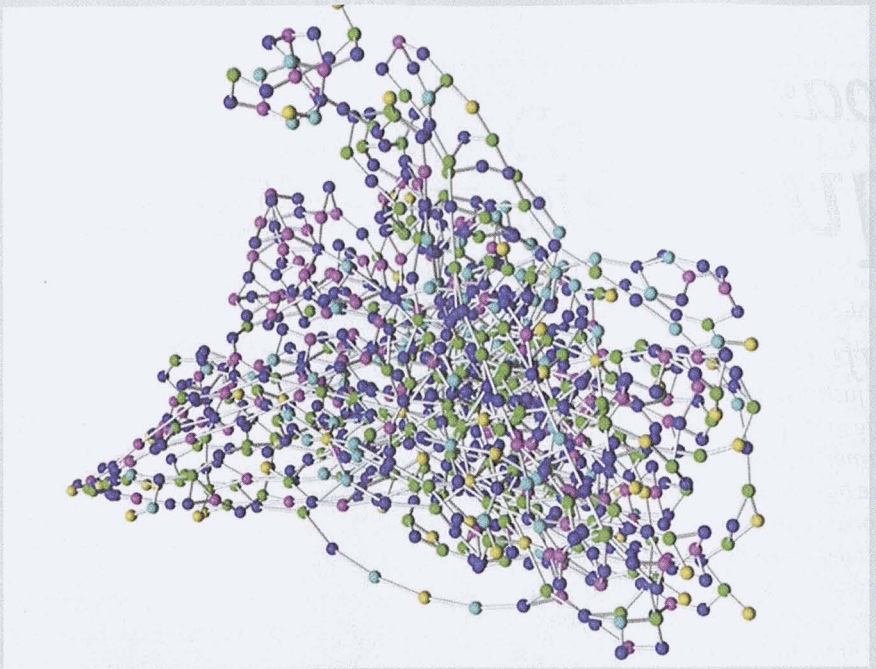
Quels types de travaux vont-ils être menés au LHS (voir l'encadré) ?

J.-Y. M. : En matière de sécurité informatique, il existe un grand nombre de travaux de recherche amont, excepté toutefois en virologie, où les travaux fondamentaux sont rares : notre équipe-projet Carte est quasiment la seule en France à travailler spécifiquement sur le sujet. Le LHS va permettre de mener des expérimentations en toute sécurité. Par exemple, l'exécution de virus, le stockage de codes malveillants, etc. sont autant de tâches qui nécessitent un milieu confiné.

Concrètement, comment se matérialise un tel laboratoire ?

J.-Y. M. : Le LHS se compose de trois salles. On pénètre dans la première avec un badge et sur la base d'une première analyse biométrique (présentation d'un doigt). Dans cette salle arrivent différents réseaux séparés tandis que les machines sont verrouillées sur place, sans aucun lien avec le réseau internet. Les salles suivantes sont les salles serveurs. Pour accéder à la troisième, celle où est stocké le matériel le plus sensible, la reconnaissance biométrique utilise l'iris de l'œil.

Au plan fonctionnel, le laboratoire se décompose en trois entités. La première a été baptisée le télescope : il s'agit d'« écouter » et de surveiller l'internet. Nous récupérons les codes malveillants, les traces d'attaques ou d'intrusions… afin de les analyser. La seconde entité est l'« éprouvette » : cette fonctionnalité consiste, soit à faire des expé-



© INRIA-J.-Y. MARION

La détection d'un virus peut se faire en prenant en compte sa structure interne, qui est représentée sous la forme d'un graphe. Celui-ci est calculé à partir d'un programme assembleur et forme une signature. L'avantage de cette signature est qu'elle résiste aux mutations locales du virus.

riences (lancer des codes malveillants, les analyser…), soit à simuler un « micro-internet » et des attaques sur ce réseau. Elle s'appuie sur une grappe de PC totalement déconnectés du monde extérieur. Enfin, la troisième unité concerne la production : nous mettons à disposition de l'extérieur (services en ligne, échanges avec des collègues…) les différents outils que nous construisons.

Cette idée de laboratoire de haute sécurité fait immanquablement penser à son analogue en biologie, d'autant que les emprunts terminologiques à cette discipline sont multiples. Vous êtres-vous inspirés des laboratoires P3 ou P4 ?

J.-Y. M. : La sécurité de l'infrastructure est différente. En revanche, la métaphore biologique existe bel et bien, mais à un autre égard. Le langage des sciences de la vie apparaît l'une des premières fois dans la première thèse connue en sécurité informatique : celle de l'informaticien américain Frederick Cohen (1986) qu'avait dirigée Léonard Adleman, coauteur de l'algorithme cryptographique RSA* et du calcul ADN. Pourquoi cette métaphore ? Dans le cas des virus, elle est presque évidente : ils se reproduisent avec parfois des mutations (en changeant leur propre code), à l'instar des virus biologiques. On peut même remonter à des travaux plus anciens, à savoir ceux de John Von Neumann (1903-1957), l'un des pionniers de l'informatique, sur les systèmes autoreproducteurs. Ce dernier avait en particulier travaillé sur la notion d'autoreproduction d'automates cellulaires.

Dans votre domaine, la virologie, quelles sont vos priorités de recherche ?

J.-Y. M. : Les aspects les plus fondamentaux portent sur la détection des codes malveillants. Au plan applicatif, l'objectif est de trouver de nouveaux algorithmes pour les détecter. Les techniques aujourd'hui mises en œuvre dans les antivirus consistent à analyser leur signature syntaxique, c'est-à-dire une suite de caractères permettant de les identifier. Le problème est que pour la moindre mutation d'un virus, il faut avoir la nouvelle signature, ce qui exclut toute possibilité d'anticipation. D'où un véritable besoin de recherche.

Mais comment savoir si un programme est, ou n'est pas, malveillant ? Pour répondre à cette question, il faut décrire ce que fait un programme, ce qui suppose de disposer d'algorithmes permettant de produire cette description. La tâche est d'autant plus ardue que les virus « se camouflent » et se modifient pendant leur exécution : on parle d'analyse de programmes auto-modifiant.

Nous explorons également une autre approche, que nous appelons analyse comportementale. L'idée est d'analyser dynamiquement la suite des actions exécutées par un programme afin de voir si cette suite d'actions doit être considérée comme malveillante, agressive, etc.

Mais ne peut-on pas imaginer un système infaillible ?

J.-Y. M. : Hélas non et, de surcroît, cela se démontre rigoureusement. On doit cette démonstration au mathématicien britannique Alan Turing (1912-1954), pionnier de l'informatique et de l'intelligence artificielle. De manière très schématique, le raisonnement est le suivant : connaître l'intention d'un programme est tout aussi difficile que de savoir si un système contient un bug. Or trouver automatiquement un bug dans un logiciel est un problème indécidable.

Alors n'y a-t-il pas moyen de mettre en place une sorte de politique de « dissuasion informatique », à l'instar de la politique de dissuasion nucléaire. Autrement dit, peut-on faire en sorte qu'un système soit suffisamment dissuasif pour éviter l'attaque ?

J.-Y. M. : *A priori* non. Toutefois, certains exemples se rapprochent de cette configuration. Par exemple, l'une des menaces actuelles sur la sécurité informatique est les *botnets* : il s'agit de réseaux reliant un grand nombre de machines « zombies » entre elles (jusqu'à plusieurs millions), le plus souvent réparties sur toute la planète. L'utilisation d'un tel réseau pour envoyer des spams, des virus, ou autres programmes malveillants est très difficile à contrer. L'idée est alors de passer à l'attaque pour neutraliser le *botnet* visé.

Typiquement, au sein du LHS, nous simulons le *botnet* dans la partie « éprouvette » du laboratoire, puis nous concevons et lançons une attaque contre le *botnet* afin de l'analyser et d'en prendre le contrôle.

Propos recueillis par Dominique Chouchan

* Travaux réalisés par l'équipe-projet Madynes que dirige Olivier Festor.

* Le sigle RSA désigne un algorithme de cryptographie parmi les plus utilisés encore à ce jour. Il est construit à partir des initiales des noms de ses trois inventeurs : Ronald Rivest, Adi Shamir et Leonard Adleman.

Le LHS en bref

Le laboratoire de haute sécurité, créé en juillet 2010 au sein du centre Inria Nancy-Grand Est, est copiloté par le Laboratoire lorrain de recherche en informatique et ses applications (Loria). Il a bénéficié de financements du Fonds européen de développement régional (Feder), de la Région Lorraine, de la Communauté urbaine du Grand Nancy et du ministère de l'Enseignement supérieur et de la Recherche. Les recherches sont menées en partenariat avec les universités lorraines, le CNRS et la Délégation générale à l'armement (DGA).

Réalisation rédactionnelle : Dominique Chouchan - Réalisation graphique : E-Graphics